ABSTRACT

Sub-daily time-scale data such as hourly base are important for the purpose of modeling the urban system. However, as similar data may not be readily available, a stochastic rainfall model is mandatory to generate reliable rainfall series that have similar properties as those of the observed in order to estimate the input for design work in the future. In this study, one of the famous models that applied the Poisson clustered point process is the Bartlett-Lewis Rectangular Pulse Model (BLRPM) will be used to access a 10-year hourly rainfall data from Station Tele Ulu Remis, Johore, Malaysia. This model applies a flexible fitting procedure to match approximately to the historical data by an optimization technique called as Shuffle Complex Evolution (SCE). SCE algorithms is chosen for the parameters estimation by minimizing an objective function with six parameters λ, κ, φ, μ_x, α and ν. The SCE algorithm performs very well in obtaining the global optimum value and the time to get the optimum value is fast. The estimated parameters for the month of November and December were compared with Powell method for validity. Subsequently, an hourly and daily rainfall simulation based on BLRPM is carried out. The performance of the BLRPM is then evaluated on a monthly basis in term of its ability to preserve statistical and physical properties. The said properties involve rainfall series over time-scales of 1-hr and 24-hr. Results from the model evaluation have indicated that BLRPM is capable to reproduce most of the statistical and physical properties of the historical data. There are also some properties that are failed to be preserved accurately. However, the BLRPM is still capable to preserve the trend of the observed properties.

ABSTRAK

Data siri masa pada skala setiap jam adalah mustahak sebagai tujuan untuk pemodelan system perbandaran. Walaubagaimanapun, disebabkan kemungkinan bahawa rekod amoun data siri setiap jam yang terhad, model hujan berstokastik adalah penting untuk menjana data sedemikian yang memiliki sifat serupa dengan data yang dicerap sebagai penilaian bagi kerja-kerja perekaan masa depan. Dalam kajian ini, salah satu model process titik berasaskan process berstokastik yang terkemuka ialah Bartlett-Lewis Rectangular Pulse Model (BLRPM) akan digunakan to menilai data siri masa setiap jam selama 10 tahun dari Stesen Tele Ulu Remis, Johor, Malaysia. Model ini menggunakan kaedah yang fleksible untuk pemadanan yang menghampiri kepada rekod amoun hujan yang sedia ada melalui teknik pengoptimuman yang dikenali sebagai Shuffle Complex Evolution (SCE). Algoritma SCE dipilih untuk menaksir nilai optimum parameter λ, κ, φ, μ_x, α and ν dengan meminimumkan tujuan fungsi. Algoritma SCE dapat memperoleh nilai optimum dengan baik dan masa yang digunakan adalah singkat. Taksiran nilai optimum parameter untuk bulanan November dan December dibanding dengan kaedah Powell sebagai pengesahan. Justeru itu, simulasi data sejam and sehari berdasarkan BLRPM akan dilakukan. Keupayaan BLRPM akan diuji secara bulanan dari segi kebolehannya untuk mengekalkan sifat cerapan. Sifat cerapan yang diuji meliputi skala masa sejam and sehari. Keputusan dari model ini mencadangkan BLRPM berupaya untuk mengekalkan kebanyakan sifat cerapan data asal. Terdapat juga beberapa sifat cerapan yang tidak dapat dikekalkan setepatnya. Walau bagaimanapun, BLRPM masih mampu mengekalkan trend yang sama dengan trend pada data asal.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS

λ	Arrival time of storm origins
κ	Dimensionless parameter
φ	Dimensionless parameter
μ_x	Average of cell depth
α	Index parameter for gamma distribution (duration of rain cell)
ν	Scale parameter for gamma distribution (duration of rain cell)
γ	Duration of each rain storm
C	number of generated rain cells
η	Duration of each rain cell
β	Arrival time of rain cell origins
$f_r(\eta)$	Probability density function for Gamma distribution
$E(\eta)$	Mean for Gamma distribution
$Var(\eta)$	Variance for Gamma distribution
Y(t)	Sum of the intensities of the individual active cells at time t.
$Y_i^{(h)}$	Aggregated total depth at time scale h
$p(h)'$	Proportion dry
μ_T	Average period of activity of a storm
μ_c	Mean cells per storm
A(k)	Autocorrelation of lag k of rainfall depths

Z	Objective function
$\overline{f}_{(h)}$	Computed statistical properties *h* based on the model expression
$f_{(h)}$	Theoretical properties *h* estimated from historical data
W_i	Weight assigned to statistical properties *h*
R_M	Root-mean-square errors
X	Intensity of rain cells
R(x)	Survival function of X
SCE-UA	Shuffle Complex Evolution – University of Arizona
S	Statistics of the historical data
S_m	Median of the simulated statistics
N(.)	Counting process
BLRPM	Bartlett-Lewis Rectangular Pulse Model
h	Level of aggregation, hour
WMO	World Meteorological Order

LIST OF APPENDICES

CHAPTER 1

INTRODUCTION

1.1 Overview

The use of the Poisson cluster-based processes in stochastic rainfall modeling had been investigated as the model has shown much potential compared to most of the traditional approaches to rainfall data generation. This is because of first and higher order of statistics can be computed with a small number of parameter inputs only. A stochastic rainfall model uses historical data to estimate the model parameters and generate a sufficiently large number of rainfall events based upon the statistical characteristics of historical rainfall data. As a result, this may be fruitful towards various engineering applications which include reservoir and sewerage design as well as floods or droughts prediction and study. The parameters will then be used to simulate the desired length of rainfall series that mimic the properties of the historical data as it has been acknowledged strongly that more adequate information could be obtained from this generated series.

Model of point rainfall time-series have the potential application in most of the hydrological problems such as the generation of the rainfall across a range of time-scales for the purpose of hydrological designs. Rainfall data in daily aggregation is used for hydrology design and applications. In conjunction to this, many stochastic models which employ daily rainfall data have been developed (e.g. Todorovic et al, 1975; Han et al, 1976; Katz, 1977; Woolhiser et al, 1982) and had been an active research topic in hydrology for many years. However, hourly data is even more important in the modeling of rainfall-runoff. Hence, a cluster-based stochastic model, namely Bartlett-Lewis Rectangular Pulse Model (BLRPM) is used to model the hourly rainfall series in this study.

1.2 Research Background

Models of point rainfall time-series have potential application to a range of hydrological problems such as the generation of rainfall across a range of time scales for hydrological design and the disaggregation of larger time-interval data for short duration application. Because of the complexity and strong dependence upon initial conditions of the precipitation process, a stochastic approach is used to a purely physical model (Smith, 1981). Besides that, a time series approach is inappropriate for small time-scale (e.g. hourly time-scale) because it may require a large number of parameters. (Pattison, 1965).

Hourly timescale data are needed for the modeling of the urban systems. However, such hourly series is not readily available as compared to the daily rainfall series. In order to overcome this problem, stochastic rainfall models are implemented to estimate input for design related works. One of the models is Bartlett-Lewis Rectangular Pulse Model (BLRPM) which uses a flexible fitting procedure to estimate a set of theoretical properties and matches them approximately to the historical sample properties. Subsequently, the theoretical properties are optimized by an optimization technique so called as Shuffle Complex Evolution - University of Arizona (SCE-UA).

Recently, the use of Poisson cluster processes in stochastic modeling of rainfall has been investigated. The cell arrivals are modeled by a Poisson cluster process, i.e. storm arrivals form a Poisson process and a cell arrival distribution is assigned to each storm; the depth and duration of the cell are modeled by exponential distribution. The cluster process may be the Bartlett-Lewis Rectangular Pulse Model (BLRPM), for which storm duration is exponentially distributed and cell arrivals are Poisson distributed as this model has been proved in reproducing the proportion of dry periods and second-order properties of the depth distribution at all time intervals (1 to 24 h) efficiently.

1.3 Problem Statements

The following show several main problems that need to be solved in this study:

1. Optimization of the objective function (1.1) to estimate optimum parameters for Bartlett-Lewis Rectangular Pulse Model (BLRPM)

$$Z = \sum_{i=1}^{m} W_i \left(1 - \frac{\overline{f}_{(h)}}{f_{(h)}}\right)^2 \text{ , m = 1, 2, 3, ..., 12} \qquad (1.1)$$

2. Hourly and daily rainfall data simulation based on the Bartlett-Lewis Rectangular Pulse Model (BLRPM)

3. Evaluating the performance of the BLPRM

1.4 Objectives of Study

The main objectives of this study are:

1. To determine the optimized values for the six parameters of BLRPM using SCE-UA algorithms

2. To simulate hourly and daily data based on the parameters

3. To compare the performance of the BLPRM in term of its ability in preserving the statistical and physical properties of the historical data graphically and quantitatively

1.5 Scope of study

In order to achieve the objective of this study, some restrictions in this study are carried out. This study will focus on the 10 years' rainfall data from Station Tele Ulu Remis, Johore with BLPRM.

There are many optimization methods to be applied in order to obtain the optimum values such as Nelder and Mead Simplex downhill Search, Simulated Annealing, Quasi Newton Search, Genetic ALgorithm and etc. However, only Shuffle Complex Evolution – University of Arizona (SCE-UA) will be focused on in this study.

In the hydrology world, there are different types of rectangular pulse models. The most famous model would be the Neyman Scott Rectangular Pulse Model (NSRPM) and the Bartlett-Lewis Rectangular Pulse Model (BLRPM). Only the randomized BLRPM will be discussed in this study.

1.6 Significance of study

This study will lead to the ability of the model to perform equally well in preserving the statistical and the physical properties. One of the potential applications from this study is to estimate the frequency of occurrence of the flooding generated from rainfall. Apart from it, the generated rainfall dataset also can be applied as input to certain distributed hydrological model in order to produce critical events' characteristics of the study area that might be otherwise not available.

1.7 Organization of the thesis

Chapter 1 will discuss about the introduction of the thesis, problem statements, objectives, scopes of the study and significance of study. Chapter 2 discusses the literature review on the overall, history and development achieved for Bartlett-Lewis Rectangular Pulse Rainfall Model (BLRPM).

It is followed by Chapter 3 that will discuss the process and formulation of BLRPM, details of the optimization method, SCE-UA for optimum parameters estimation and procedure for hourly and daily rainfall simulation. On the other hand, Chapter 4 discusses results of the historical and fitted properties of the optimization method and the simulation of the rainfall data as well as its physical and statistical properties. Some comparisons and conclusion are made.

Lastly, discussions on the conclusion of the study as well as suggestions for future research are enclosed in the Chapter 5.